

Datenbanken und SQL

Kapitel 3

Datenbankdesign – Teil I: Normalformen

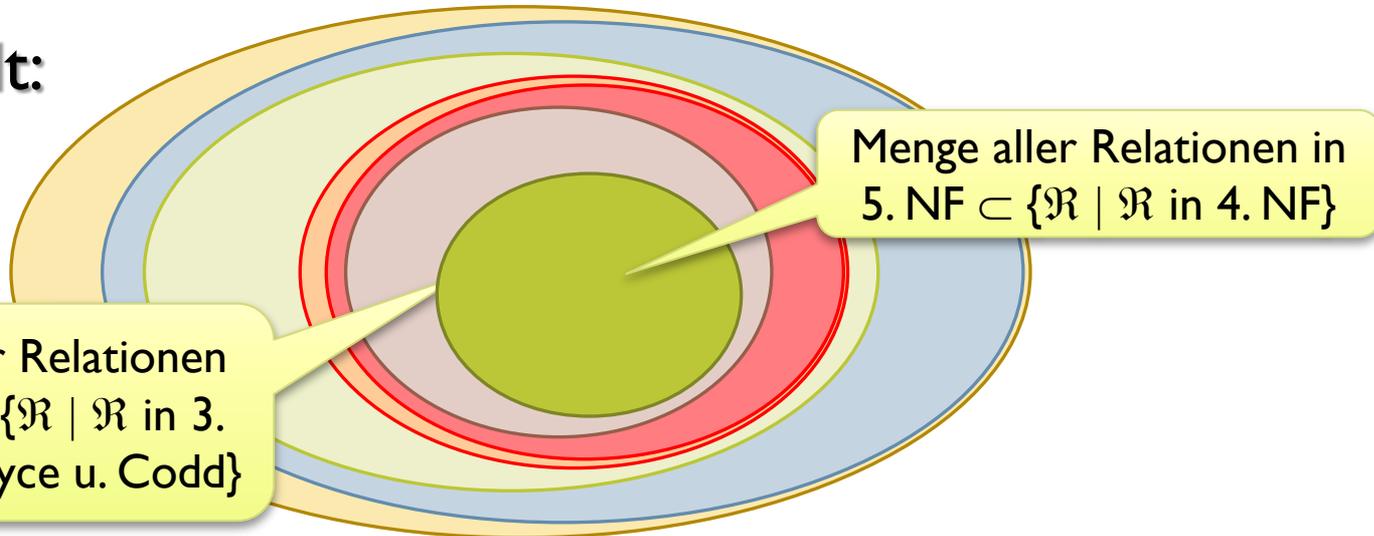
Datenbankdesign

▶ Normalformen

- ▶ *1. Normalform*
- ▶ *Funktionale Abhängigkeit*
- ▶ *2. Normalform*
- ▶ *3. Normalform nach Boyce und Codd*
- ▶ *3. Normalform nach Codd*
- ▶ *Mehrwertige Abhängigkeit und 4. Normalform*
- ▶ *Verbundabhängigkeit und 5. Normalform*

Die Normalformen im Überblick

- ▶ Es gibt 6 Definitionen von Normalformen
 - ▶ Die 1. Normalform (NF) schränkt am wenigsten ein
 - ▶ Es gibt 2 Definitionen für die 3. NF
 - ▶ Die 5. NF schränkt am stärksten ein
- ▶ Es gilt:



Menge aller Relationen
in 4. NF $\subset \{\mathcal{R} \mid \mathcal{R} \text{ in 3. NF nach Boyce u. Codd}\}$

Menge aller Relationen in
5. NF $\subset \{\mathcal{R} \mid \mathcal{R} \text{ in 4. NF}\}$

Für die Normalformen gilt:

▶ { 1. NF }

▶ ⊃ { 2. NF }

Wichtig!

▶ ⊃ { 3. NF }

Ganz wichtig!

▶ ⊃ { 3. NF BC }

Für Interessierte

▶ ⊃ { 4. NF }

▶ ⊃ { 5. NF }

▶ 3. NF = 3. NF nach Codd

3. NF BC = 3. NF nach Boyce und Codd

Definition (1. Normalform)

- ▶ Eine Relation ist in erster Normalform (1. NF), wenn alle zugrundeliegende Gebiete nur atomare Werte enthalten.
- ▶ **Folgerung:**
 - ▶ Jede (normalisierte) Relation ist in 1. NF
- ▶ **Die 1. NF ist historisch bedingt:**
 - ▶ In der Originaldefinition von Relationen war die Atomarität nicht gefordert
- ▶ **NF² = NFNF = NonFirstNormalForm**
 - ▶ Relationen, die nicht in 1. NF sind
 - ▶ NF² ist Basis für objektrelationale Datenbanken

Beispiel: Relation VerkäuferProdukt

VerkNr	VerkName	PLZ	VerkAdresse	Produktname	Umsatz
V1	Meier	80075	München	Waschmaschine	11000
V1	Meier	80075	München	Herd	5000
V1	Meier	80075	München	Kühlschrank	1000
V2	Schneider	70038	Stuttgart	Herd	4000
V2	Schneider	70038	Stuttgart	Kühlschrank	3000
V3	Müller	50083	Köln	Staubsauger	1000

- ✓ VerkäuferProdukt enthält nur atomare Werte
- ✓ VerkäuferProdukt ist eine Relation
- ✓ VerkäuferProdukt ist in 1. NF

VerkäuferProdukt (Wiederholung)

▶ Redundanz

- ▶ Je mehr ein Verkäufer verkauft, um so häufiger in Tabelle!
- ▶ Ändert sich die Adresse eines Verkäufers, muss dies in allen entsprechenden Einträgen erfolgen. Sonst: Inkonsistenz!

▶ Handhabung

- ▶ Soll Produkt Staubsauger aus dem Sortiment genommen werden, so ist auch Verkäufer Müller zu löschen!?
- ▶ Verkäufer Schmidt kann erst eingetragen werden, wenn er etwas verkauft hat!?

Warum verwenden wir Tabellen?

▶ Beispiel: Tabelle der Länder (Stand 2012)

Land	Einwohner	Fläche	Kennzeichen	Hauptstadt
Deutschland	81.800.000	357.121	D	Berlin
Frankreich	63.600.000	551.500	F	Paris
Großbritannien	63.200.000	243.610	UK	London
Italien	60.900.000	301.340	I	Rom
Niederlande	16.700.000	61.543	NL	Amsterdam
Polen	38.200.000	312.685	PL	Warschau
Spanien	46.200.000	505.370	E	Madrid



- ▶ Alle Einträge hängen eindeutig von der Spalte Land ab:
 - ▶ Aus dem Land folgen eindeutig alle anderen Spalten!

Definition (Funktionale Abhängigkeit)

- ▶ Ein Attribut Y einer Relation R heißt funktional abhängig vom Attribut X derselben Relation, wenn zu jedem X -Wert höchstens ein Y -Wert möglich ist.
- ▶ Informell: „Aus X folgt eindeutig Y “
- ▶ Wir schreiben: $X \rightarrow Y$

- ▶ Land \rightarrow Einwohner
- ▶ Land \rightarrow Fläche
- ▶ Land \rightarrow Kennzeichen
- ▶ Land \rightarrow Hauptstadt

Folgerungen zur funktionalen Abh.

- ▶ **Primärschlüssel** → alle anderen Attribute
 - ▶ (da Primärschlüssel eindeutig jedes Tupel identifizieren)
- ▶ **Schlüsselkandidat** → alle anderen Attribute
 - ▶ (da Schlüsselkandidaten eindeutig jedes Tupel identifizieren)
- ▶ **Superschlüssel** → alle anderen Attribute
 - ▶ (da Superschlüssel eindeutig jedes Tupel identifizieren)

VerkäuferProdukt: Funktionale Abh.

VerkNr	VerkName	PLZ	VerkAdresse	Produktname	Umsatz
V1	Meier	80075	München	Waschmaschine	11000
V1	Meier	80075	München	Herd	5000
V1	Meier	80075	München	Kühlschrank	1000
V2	Schneider	70038	Stuttgart	Herd	4000
V2	Schneider	70038	Stuttgart	Kühlschrank	3000
V3	Müller	50083	Köln	Staubsauger	1000

- ✓ VerkNr → VerkName
- ✓ VerkNr → PLZ
- ✓ VerkNr → VerkAdresse

- ✓ (VerkNr, Produktname) → Umsatz
- ✓ PLZ → VerkAdresse ?

Problem: Abhängigkeit PLZ und Adresse

▶ Folgt aus PLZ der Ort?

- ▶ Definition Ort: Alle selbstständigen Gemeinden Deutschlands
 - ▶ Antwort: NEIN, da einige kleine Gemeinden gleiche PLZ
- ▶ Definition Ort: Alle Gemeinden mit mehr als 20000 Einwohner
 - ▶ Antwort: JA, da alle großen Gemeinden unterschiedliche PLZ

▶ Folgt aus Adresse (Ort+Straße+Nr) die PLZ?

- ▶ Antwort: NEIN
- ▶ Beispiel: Es gibt mehrere Neustadt mit Bahnhofstraße I
- ▶ Dies war mit ein Grund zur Einführung der PLZ!

▶ Folgt aus der PLZ das Bundesland?

- ▶ Antwort: JA, da bei der Einführung der PLZ darauf Rücksicht genommen wurde

Problem: Fehlende Minimalität

▶ Schlüsselkandidat → alle anderen Attribute

- ▶ Schlüsselkandidaten sind minimal
- ▶ Jeder Wert kommt nur einmal vor

▶ Superschlüssel → alle anderen Attribute

- ▶ Superschlüssel sind nicht notwendigerweise minimal! Problem!

▶ Beispiel:

- ▶ (Verknr, Produktname) → Umsatz Primärschlüssel!
- ▶ (Verknr, VerkName, Produktname) → Umsatz Superschlüssel
- ▶ (Verknr, PLZ, Produktname) → Umsatz Superschlüssel, usw.

▶ Folgerung:

- ▶ Invasion weiterer wertloser funktionaler Abhängigkeiten

Definition (Volle funktionale Abh.)

- ▶ Ein Attribut Y einer Relation R heißt voll funktional abhängig vom Attribut X derselben Relation, wenn
 - ▶ es funktional abhängig ist von X
 - ▶ es nicht funktional abhängig ist von beliebigen Teilattributen von X
- ▶ Wir schreiben: $X \Rightarrow Y$
- ▶ Folgerung:
 - ▶ Es gilt immer: Primärschlüssel \rightarrow alle anderen Attribute
 - ▶ Es gilt nicht immer: Primärschlüssel \Rightarrow alle anderen Attribute

VerkäuferProdukt: Volle funkt. Abh.

VerkNr	VerkName	PLZ	VerkAdresse	Produktname	Umsatz
V1	Meier	80075	München	Waschmaschine	11000
V1	Meier	80075	München	Herd	5000
V1	Meier	80075	München	Kühlschrank	1000
V2	Schneider	70038	Stuttgart	Herd	4000
V2	Schneider	70038	Stuttgart	Kühlschrank	3000
V3	Müller	50083	Köln	Staubsauger	1000

- ✓ VerkNr \Rightarrow VerkName
- ✓ VerkNr \Rightarrow PLZ
- ✓ VerkNr \Rightarrow VerkAdresse
- ✓ (VerkNr, Produktname) \Rightarrow Umsatz

Es gibt keine weiteren vollen funktionalen Abhängigkeiten!

Eventuell PLZ \Rightarrow VerkAdresse

Definition (Zweite Normalform)

- ▶ Eine Relation ist in der zweiten Normalform (2. NF), wenn sie in der ersten Normalform ist, und jedes Nichtschlüsselattribut voll funktional vom Primärschlüssel abhängt.
- ▶ **Bemerkungen:**
 - ▶ Die 2. NF bezieht sich nur auf Primärschlüssel, nicht auf alternative Schlüssel
 - ▶ Die Relation VerkäuferProdukt ist nicht in der 2. NF

Wichtige Folgerungen zur 2. NF

▶ 1. NF:

- ▶ Primärschlüssel \rightarrow alle Nichtschlüsselattribute

▶ 2. NF:

- ▶ Primärschlüssel \Rightarrow alle Nichtschlüsselattribute

▶ **Primärschlüssel ist ein einzelnes Attribut:**

- ▶ Dann folgt: Relation ist in mindestens 2. NF

▶ **Jede Relation in 1. NF lässt sich in die 2. NF überführen:**

- ▶ Hinzufügen eines Zählers als Primärschlüssel (einzelnes Attr.)

VerkäuferProdukt2NF

Nr	VerkNr	VerkName	PLZ	VerkAdresse	Produktname	Umsatz
1	V1	Meier	80075	München	Waschmaschine	11000
2	V1	Meier	80075	München	Herd	5000
3	V1	Meier	80075	München	Kühlschrank	1000
4	V2	Schneider	70038	Stuttgart	Herd	4000
5	V2	Schneider	70038	Stuttgart	Kühlschrank	3000
6	V3	Müller	50083	Köln	Staubsauger	1000

- ▶ Primärschlüssel: Nr also: 2. NF
- ▶ Noch mehr Redundanzen (Attribut Nr)
- ▶ Aber: Weniger Anomalien (Entfernen von Staubsauger!)

VerkäuferProdukt2NF

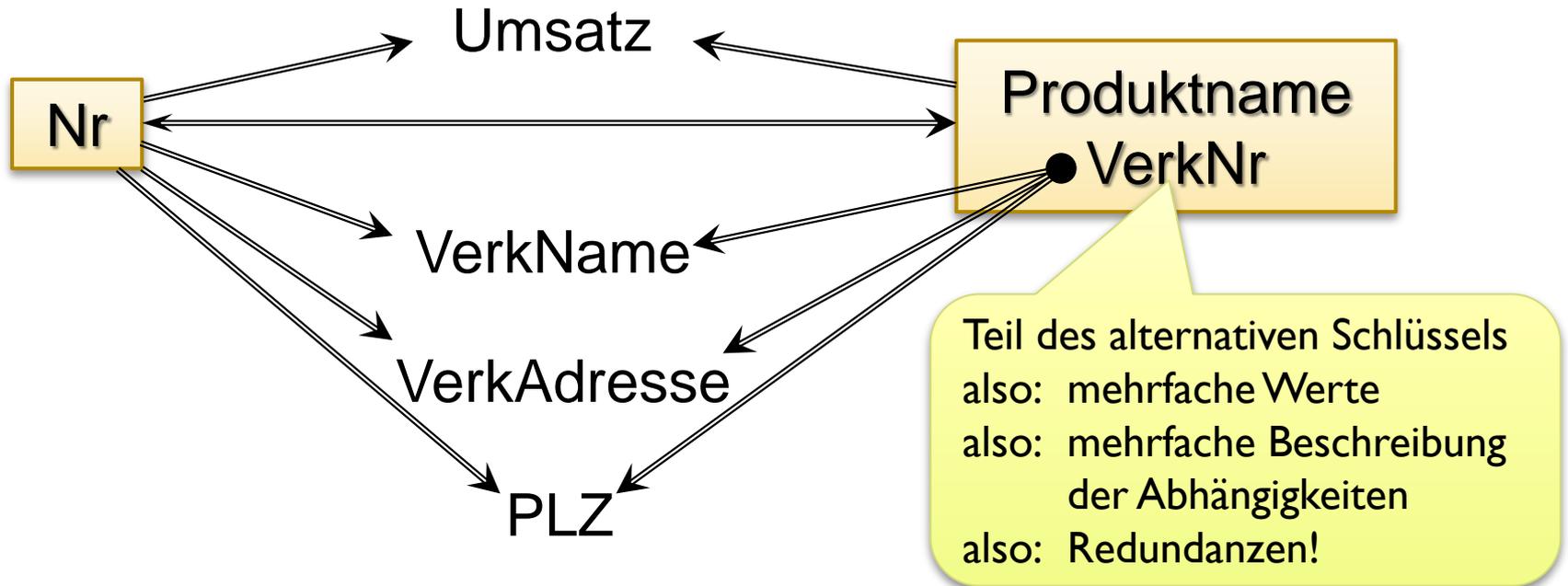
Nr	VerkNr	VerkName	PLZ	VerkAdresse	Produktname	Umsatz
1	V1	Meier	80075	München	Waschmaschine	11000
2	V1	Meier	80075	München	Herd	5000
3	V1	Meier	80075	München	Kühlschrank	1000
4	V2	Schneider	70038	Stuttgart	Herd	4000
5	V2	Schneider	70038	Stuttgart	Kühlschrank	3000
6	V3	Müller	50083	Köln	NULL	1000

- ▶ Primärschlüssel: Nr also: 2. NF
- ▶ Noch mehr Redundanzen (Attribut Nr)
- ▶ Aber: Weniger Anomalien (Entfernen von Staubsauger!)

jetzt erlaubt

VerkaeufersProdukt2NF

- ▶ Volle funktionale Abhängigkeiten:



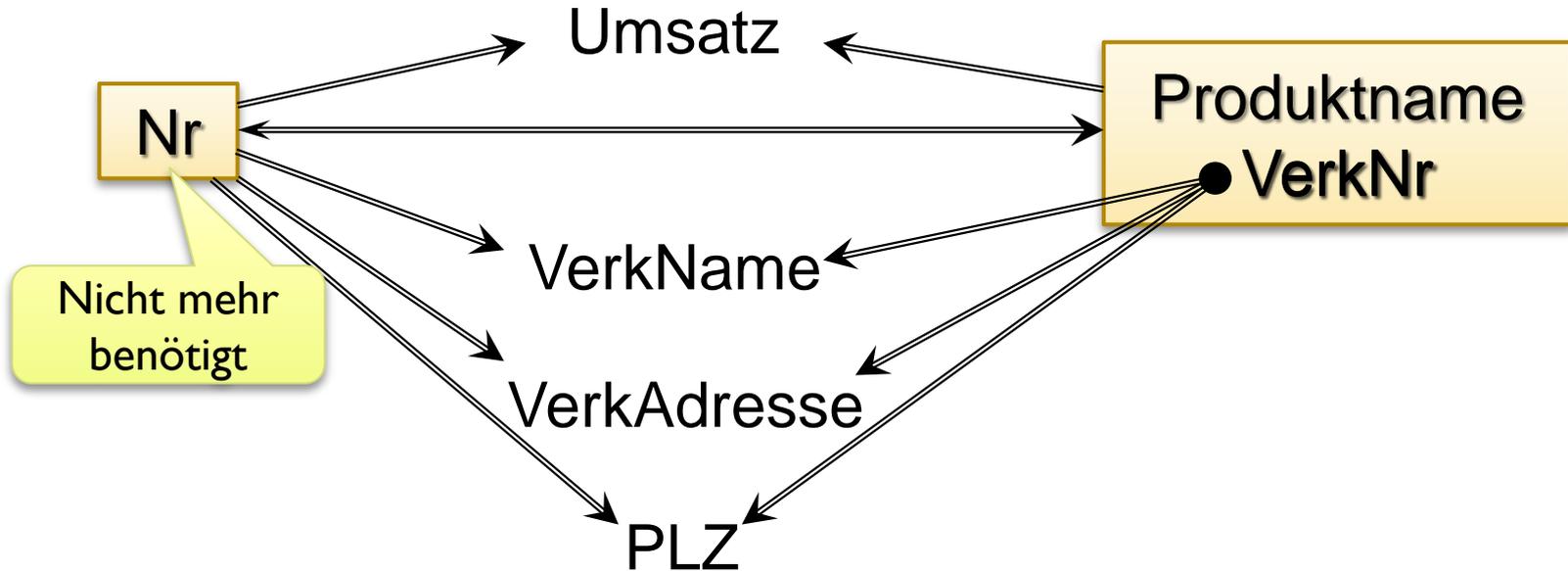
Definition (Determinante)

- ▶ Eine Determinante ist ein (eventuell zusammengesetztes) Attribut, von dem ein anderes voll funktional abhängt.
- ▶ **Bemerkungen:**
 - ▶ Ein wertvolles Hilfsmittel
 - ▶ Alle Attribute, von denen Doppelpfeile ausgehen, sind Determinanten
 - ▶ Determinanten in VerkäuferProdukt2NF:
 - ▶ Nr,Verknr, (Verknr, Produktname)

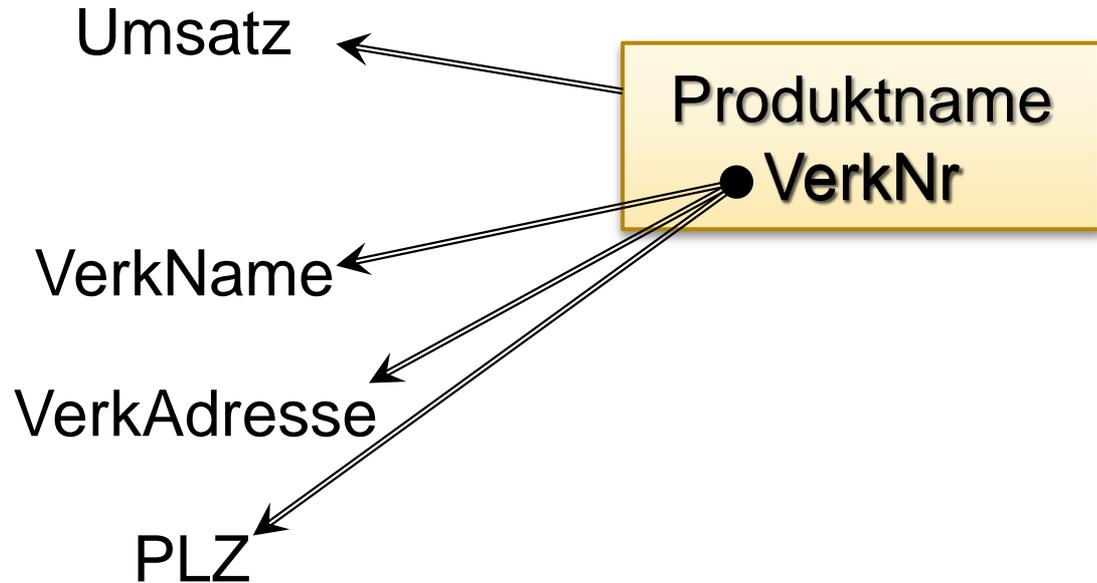
Dritte Normalform nach Boyce u. Codd

- ▶ Eine normalisierte Relation ist in der dritten Normalform, wenn jede Determinante dieser Relation ein Schlüsselkandidat ist.
- ▶ **Bemerkungen:**
 - ▶ Alle Abhängigkeiten von nicht eindeutigen Werten (Schlüsselkandidaten) werden verboten!
 - ▶ Damit ist eine Relation in 3. NF redundanzfrei (außerhalb der Schlüsselkandidaten)
 - ▶ VerkäuferProdukt2NF ist nicht in 3. NF
 - ▶ denn: VerkNr ist kein Schlüsselkandidat
 - ▶ Jede Relation in 2. NF lässt sich in Relationen der 3. NF überführen

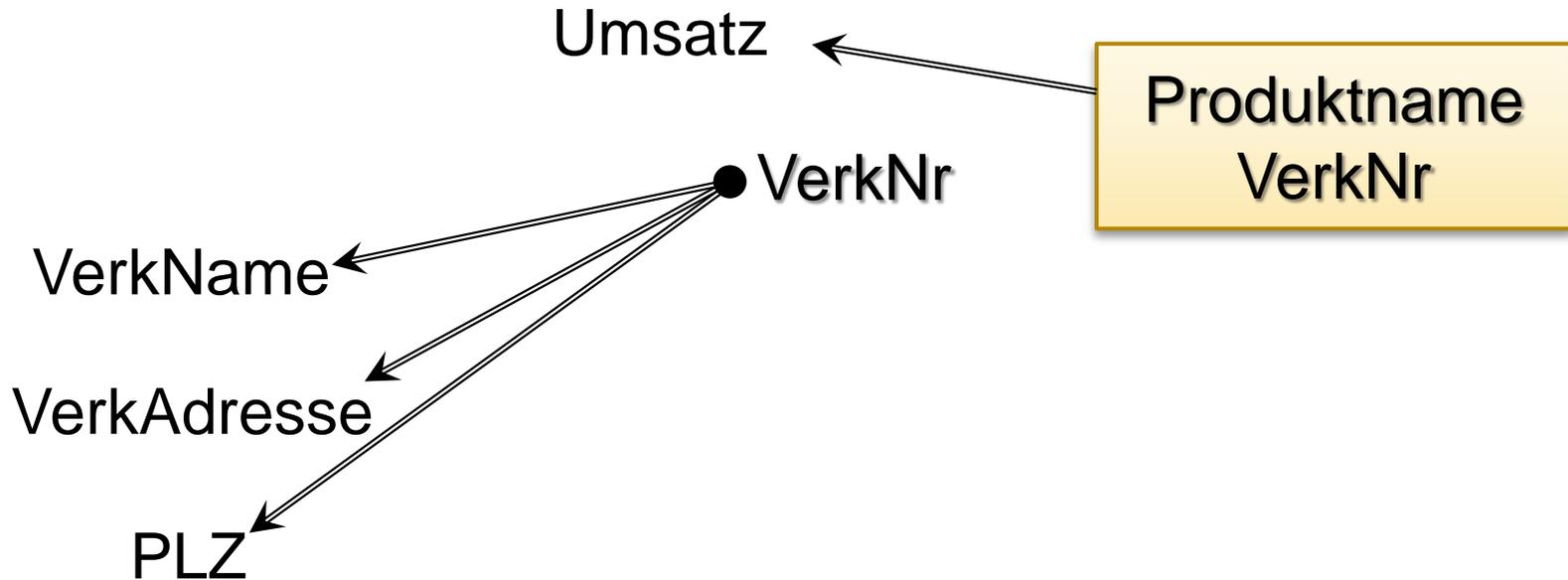
Überführung in die 3. NF



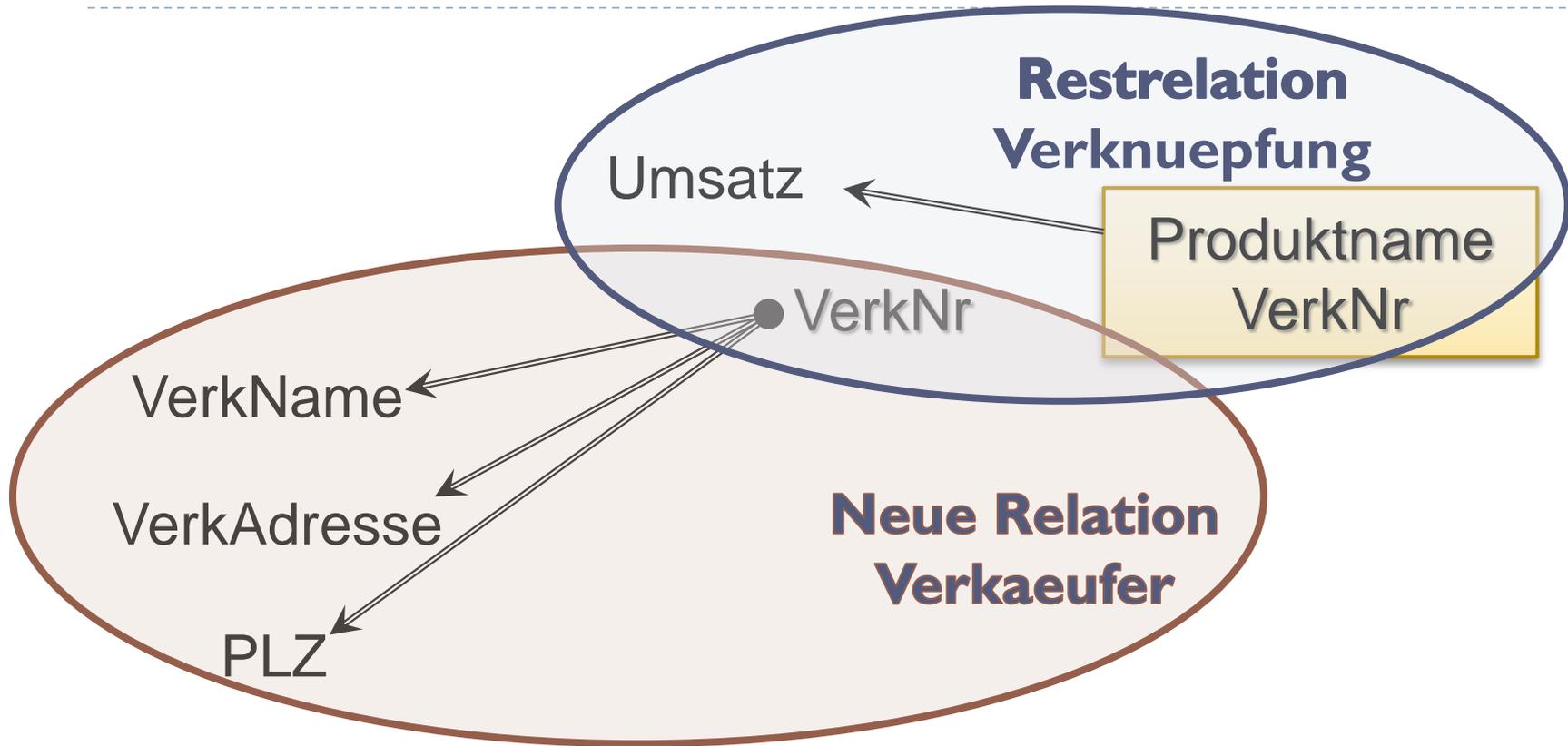
Überführung in die 3. NF



Überführung in die 3. NF



Überführung in die 3. NF



Verkäufer und Produkte (Schritt 1)

Restrelation Verknüpfung:

VerkNr	Produktname	Umsatz
V1	Waschmaschine	11000
V1	Herd	5000
V1	Kühlschrank	1000
V2	Herd	4000
V2	Kühlschrank	3000
V3	Staubsauger	1000

Neue Relation Verkäufer:

VerkNr	VerkName	PLZ	VerkAdresse
V1	Meier	80075	München
V2	Schneider	70038	Stuttgart
V3	Müller	50083	Köln

Alles ist in Ordnung, da 3. NF; aber:

Wir haben den Verkäufer herausgenommen, warum nicht auch das Produkt?

Verkaefer und Produkte (Schritt 2)

Relation Verkaefer:

VerkNr	VerkName	PLZ	VerkAdresse
V1	Meier	80075	München
V2	Schneider	70038	Stuttgart
V3	Müller	50083	Köln

Relation
Produkt:

ProdNr	Produktname
P1	Waschmaschine
P2	Herd
P3	Kühlschrank
P4	Staubsauger

Relation Verknuepfung:

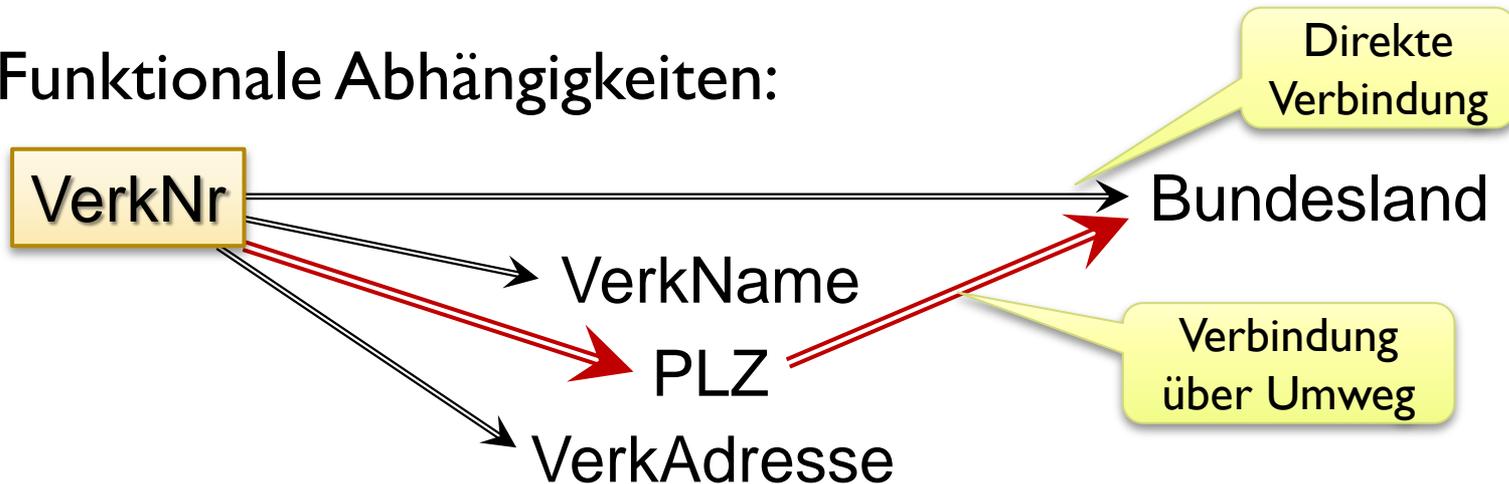
VerkNr	ProdNr	Umsatz
V1	P1	11000
V1	P2	5000
V1	P3	1000
V2	P2	4000
V2	P3	3000
V3	P4	1000

Weitere Produkteigen-
schaften jetzt möglich!

VerkaeuferLand

VerkNr	VerkName	PLZ	VerkAdresse	Bundesland
V1	Meier	80075	München	Bayern
V2	Schneider	70038	Stuttgart	Baden-Württemberg
V3	Müller	50083	Köln	Nordrhein-Westfalen

► Funktionale Abhängigkeiten:



Transitive Abhängigkeit

- ▶ Ein Attribut Y einer Relation R heißt transitiv abhängig vom Attribut X derselben Relation, wenn ein Nichtschlüssel-Attribut Z existiert, so dass gilt:
 - Das Attribut Z hängt voll funktional vom Attribut X und das Attribut Y voll funktional vom Attribut Z ab.
- wenn also ein Z existiert mit: $X \Rightarrow Z \Rightarrow Y$
- ▶ VerkäuferLand: $\text{VerkNr} \Rightarrow \text{PLZ} \Rightarrow \text{Bundesland}$
- ▶ Transitive Abhängigkeit des Bundeslands von der Verkäufernummer!

Dritte Normalform nach Codd

- ▶ Eine Relation ist in der dritten Normalform (nach Codd), wenn sie sich in der zweiten Normalform befindet und jedes Nichtschlüsselattribut nicht transitiv vom Primärschlüssel abhängt.
- ▶ Relation VerkäuferLand ist nicht in der 3. NF nach Codd
- ▶ VerkäuferLand ist nicht in der 3. NF nach Boyce u. Codd
 - ▶ denn: PLZ ist Determinante, aber kein Schlüsselkandidat
- ▶ Gibt es einen Unterschied zwischen den beiden 3. NFs?

Vergleich der dritten Normalformen

- ▶ **3. NF nach Boyce und Codd → 3. NF nach Codd**
 - ▶ Beweis:
 - ▶ Transitive Abhängigkeiten bedingen eine Determinante, die nicht Schlüsselkandidat ist.
 - ▶ 3. NF nach Boyce u. Codd → keine solchen Determinanten → keine transitiven Abhängigkeiten → 3. NF nach Codd
- ▶ **Das Umgekehrte gilt nicht!**
 - ▶ Es gibt Relationen in 3. NF nach Codd, die nicht in 3. NF nach Boyce u. Codd sind.
 - ▶ Übung! Bitte die nächste Folie beachten!

Hinweise zur dritten Normalform

- ▶ Für Relationen R mit einfachen (nicht zusammengesetzten) Schlüsselkandidaten gilt:
 - ▶ Beide Definitionen der 3. NF sind gleichwertig
- ▶ Es gilt (ohne Beweis):
 - ▶ Unterschiede kann es nur dann geben, wenn zwei zusammengesetzte Schlüsselkandidaten existieren, die ein gemeinsames Attribut besitzen
- ▶ Die dritten Normalformen beseitigen alle Redundanzen und Anomalien außerhalb der Schlüsselkandidaten
- ▶ Also: Wir definieren nur sinnvolle Schlüsselkandidaten!

VerkäuferProdukt3NF

VerkName	PLZ	VerkAdresse	Produktname	Umsatz
Meier	80075	München	Waschmaschine	11000
Meier	80075	München	Herd	5000
Meier	80075	München	Kühlschrank	1000
Schneider	70038	Stuttgart	Herd	4000
Schneider	70038	Stuttgart	Kühlschrank	3000
Müller	50083	Köln	Staubsauger	1000

- ▶ **Voraussetzung:**
 - ▶ (VerkName, PLZ, VerkAdresse) identifiziert Verkäufer eindeutig
 - ▶ Gegebenenfalls Zusätze: komplette Adresse, Vorname, junior usw.
- ▶ **Nur ein Schlüsselkandidat und eine Determinante:**
 - ▶ (VerkName, PLZ, VerkAdresse, Produktname)

Problem der dritten Normalform

- ▶ **3. NF beseitigt keine Redundanzen innerhalb der Schlüsselkandidaten**
- ▶ **Lösung**
 - ▶ Wir erstellen Relationen mit vernünftigen Schlüsselkandidaten
- ▶ **Was ist vernünftig?**
 - ▶ Gesunden Menschenverstand anwenden, oder:
 - ▶ Weitere Normalformen studieren
- ▶ **Achtung:**
 - ▶ Relation **VerkaeuerProdukt3NF** ist in 4. NF (siehe später)!

Relation VerkäuferProduktKFZ

VerkNr	Produktname	KFZNr
1	Waschmaschine	M-E 515
1	Waschmaschine	M-X 333
1	Herd	M-E 515
1	Herd	M-X 333
1	Kühlschrank	M-E 515
1	Kühlschrank	M-X 333
2	Herd	S-H 654
2	Herd	K-J 123
2	Kühlschrank	S-H 654
2	Kühlschrank	K-J 123
3	Staubsauger	K-J 123

▶ Es gelte:

- ▶ Meier (VerkNr 1) benutzt M-E 515 und M-X 333
- ▶ Schneider (Nr 2) benutzt S-H 654 und K-J 123
- ▶ Müller (VerkNr 3) benutzt K-J 123

▶ Keine funktionalen Abh.

▶ Primärschlüssel:

- ▶ (VerkNr, Produktname, KFZNr)

Probleme von VerkäuferProduktKFZ

- ▶ **Relation ist in 3. NF**
- ▶ **Primärschlüssel enthält aber Redundanzen**
- ▶ **Anomalien treten auf:**
 - ▶ Die erste Zeile kann nicht gelöscht werden, ohne dass auch andere Zeile gelöscht werden müssen
 - ▶ Verkauft Verkäufer Schneider (Verknr 2) Staubsauger, so müssen 2 Zeilen eingefügt werden
- ▶ **Problem:**
 - ▶ Es wurden 2 funktionale Abhängigkeiten ineinander gemengt, die voneinander aber völlig unabhängig sind:
 - ▶ Verkäufer verkauft Produkte und Verkäufer fährt mit KFZ

Mehrwertige Abhängigkeit

- ▶ Ein Attribut Y einer Relation ist von einem Attribut X dieser Relation mehrwertig abhängig ($X \twoheadrightarrow Y$), wenn ein weiteres Attribut Z dieser Relation existiert mit den Eigenschaften:
 - Ein Y -Attributwert hängt vom dazugehörigen (X,Z) -Paar bereits allein eindeutig vom X -Wert ab und ist unabhängig vom Z -Attribut.
 - Das Attribut X ist minimal.
- ▶ Es gilt in `VerkaeuerProduktKFZ`
 - ▶ $X = \text{Verknr}$, $Y = \text{Produktname}$, $Z = \text{KFZNr}$,
 - ▶ also: $\text{Verknr} \twoheadrightarrow \text{Produktname}$

Infos zur mehrwertigen Abhängigkeit

- ▶ Aus $X \twoheadrightarrow Y$ (über Z) folgt $X \twoheadrightarrow Z$ (über Y)
 - ▶ Begründung: Symmetrie zwischen den beiden Abhängigkeiten
- ▶ Also:
 - ▶ Verknr \twoheadrightarrow Produktname
 - ▶ Verknr \twoheadrightarrow KFZNr
- ▶ Aus $X \Rightarrow Y$ folgt $X \twoheadrightarrow Y$
- ▶ Die mehrwertige Abhängigkeit ist eine Verallgemeinerung der (vollen) funktionalen Abhängigkeit
 - ▶ Beweis: Setzen wir in der Definition der mehrwertigen Abh. $Z = \emptyset$, so ist die funktionale Abh. gegeben

Überführung in NF²

- ▶ Verletzen wir die Atomarität, so lässt sich Relation **VerkaeuerProduktKFZ** einfach abbilden:

VerkNr	Produktname	KFZNr
1	Waschmaschine	M-E 515
	Herd	M-X 333
	Kühlschrank	
2	Herd	S-H 654
	Kühlschrank	K-J 123
3	Staubsauger	K-J 123

- ▶ **Hinweis:**
 - ▶ Wir verbieten solche Relationen in der 4. NF
 - ▶ Die nicht atomare Relation zeigt, wie wir zerlegen können!

Vierte Normalform

- ▶ Eine normalisierte Relation ist in der vierten Normalform, wenn aus jeder mehrwertigen Abhängigkeit $X \twoheadrightarrow Y$ folgt, dass X ein Schlüsselkandidat ist.
- ▶ VerkäuferProduktKFZ besitzt zwei mehrwertige Abhängigkeiten und VerkNr ist kein Schlüsselkandidat.
 - ▶ Also: Keine 4. NF
- ▶ Jede Relation mit mehrwertigen Abhängigkeiten lässt sich in seine Abhängigkeiten zerlegen, bei zwei mehrwertigen Abhängigkeiten also in zwei Relationen!

Zerlegung von VerkäuferProduktKFZ

Relation VerkäuferProduktname:

VerkNr	Produktname
1	Waschmaschine
1	Herd
1	Kühlschrank
2	Herd
2	Kühlschrank
3	Staubsauger

Relation VerkäuferKFZ:

VerkNr	KFZNr
1	M-E 515
1	M-X 333
2	S-H 654
2	K-J 123
3	K-J 123

- ▶ Beide Relationen sind in 4. NF und optimal
- ▶ Bei vielen Verkäufen und Verwendung vieler KFZ haben beide Relationen wesentlich weniger Redundanz als die Originalrelation

Weitere Probleme mit 4. Normalform

- ▶ **Gewünscht:** Angabe der Kilometer, die Verkäufer mit KFZ gefahren ist, abhängig vom verkauften Produkt und Jahr:

VerkNr	Produktname	KFZNr	Jahr	KM
I	Waschmaschine	M-E 515	2011	622
I	Waschmaschine	M-E 515	2012	1105
I	Waschmaschine	M-X 333	2011	305
I	Waschmaschine	M-X 333	2012	0
I	Herd	M-E 515	2011	912
I	Herd	M-E 515	2012	1111
I	Herd	M-X 333	2011	0
I	Herd	M-X 333	2012	222
I	Kühlschrank	M-E 515	2011	333
...

Fragen zu dieser Relation

- ▶ Schlüsselkandidaten, Primärschlüssel?
- ▶ Volle funktionale Abhängigkeiten?
- ▶ Mehrwertige Abhängigkeiten?
- ▶ Normalform?
- ▶ Wenn nicht 4. NF, wie zerlegen wir diese Relation?

- ▶ Antwort:
 - ▶ Übung

VerkäuferProdukt4NF

Tupel nicht einzeln lösbar

VerkNr	Produktname	KFZNr
1	Waschmaschine	M-E 515
1	Herd	M-E 515
1	Herd	M-X 333
1	Kühlschrank	M-E 515
2	Herd	S-H 654
2	Herd	K-J 123
2	Kühlschrank	S-H 654
3	Staubsauger	K-J 123

Löschen möglich!

Annahme:

- Spezialhalterung für Waschmaschinen und Kühlschränke gibt es nicht in:
 - ▶ M-X 333 und K-J 123

- ▶ **Produktname und KFZNr hängen jetzt voneinander ab**
- ▶ **Es gibt keine mehrwertigen Abhängigkeiten, also: 4. NF**
- ▶ **Aber: Es treten immer noch Anomalien auf:**
 - ▶ **Beispiel: Nicht jeder Herd lässt sich einzeln löschen!!**

VerkaeufersProdukt4NF

- ▶ **Es gibt auch Einfügeanomalien!**
- ▶ **Ein Zerlegen in 2 Relationen ist nicht möglich ohne Verlust an Information!**
 - ▶ Zerlegen in `VerkaeufersProduktname` und `VerkaeufersKFZ` bringt nichts, da
$$\text{VerkaeufersProduktname} \bowtie \text{VerkaeufersKFZ} = \text{VerkaeufersProduktKFZ} \quad !!!$$
 - ▶ Die Info über die fehlenden Halterungen in einigen KFZ geht verloren!
- ▶ **Aber: Zerlegen mit anschließendem Verbund ist ein guter Ansatz**

Definition (Verbundabhängigkeit)

- ▶ Eine Relation R besitzt eine **Verbundabhängigkeit**, wenn sie mittels Projektion nicht trivial in Teilrelationen zerlegt werden kann, so dass der Verbund dieser Teilrelationen wieder die Relation R ergibt.
- ▶ **Nicht trivial** heißt, dass die Teilrelationen jeweils unterschiedliche Primärschlüssel besitzen.
- ▶ **Beispiel. Mit**
 - ▶ $R = \pi_{\text{Projektion1}}(R) \bowtie \pi_{\text{Projektion2}}(R) \bowtie \pi_{\text{Projektion3}}(R)$
 - ▶ Projektion 1 bis 3 sind nicht trivial
 - ▶ Dann besitzt R eine **Verbundabhängigkeit**

Verbundabhängigkeit

- ▶ Relation `VerkaeuerProduktKFZ` besitzt nicht nur eine mehrwertige Abhängigkeit, sondern auch eine Verbundabhängigkeit wegen

$\text{VerkaeuerProduktname} \bowtie \text{VerkaeuerKFZ} = \text{VerkaeuerProduktKFZ}$
 $\text{VerkaeuerProduktname} = \pi_{\text{Verknr,Produktname}}(\text{VerkaeuerProduktKFZ})$
 $\text{VerkaeuerKFZ} = \pi_{\text{Verknr,KFZNr}}(\text{VerkaeuerProduktKFZ})$

- ▶ Dies gilt allgemein (ohne Beweis):
 - ▶ Aus funktionaler Abhängigkeit folgt mehrwertige Abhängigkeit
 - ▶ Aus mehrwertiger Abhängigkeit folgt Verbundabhängigkeit

Definition (Fünfte Normalform)

- ▶ Eine Relation ist in der fünften Normalform, wenn sie in der vierten Normalform ist und keine Verbundabhängigkeiten besitzt.
- ▶ In der Praxis:
 - ▶ Es ist extrem schwer, Verbundabhängigkeiten zu finden
 - ▶ Diese Abhängigkeiten müssen noch nicht vorhanden sein und könnten erst in Zukunft auftreten!
 - ▶ Die 5. NF spielt daher fast keine Rolle
- ▶ Ist Relation VerkäuferProdukt4NF in der 5. NF?

Zerlegung von VerkäuferProdukt4NF

$\text{VerkäuferProduktname} = \pi_{\text{VerkNr, Produktname}}(\text{VerkäuferProdukt4NF})$

$\text{VerkäuferKFZ} = \pi_{\text{VerkNr, KFZNr}}(\text{VerkäuferProdukt4NF})$

$\text{ProduktKFZ} = \pi_{\text{Produktname, KFZNr}}(\text{VerkäuferProdukt4NF})$

VerkäuferProduktname

VerkNr	Produktname
1	Waschmaschine
1	Herd
1	Kühlschrank
2	Herd
2	Kühlschrank
3	Staubsauger

VerkäuferKFZ

VerkNr	KFZNr
1	M-E 515
1	M-X 333
2	S-H 654
2	K-J 123
3	K-J 123

ProduktKFZ

Produktname	KFZNr
Waschmaschine	M-E 515
Herd	M-E 515
Herd	M-X 333
Herd	S-H 654
Herd	K-J 123
Kühlschrank	M-E 515
Kühlschrank	S-H 654
Staubsauger	K-J 123

  \bowtie : VerkäuferProduktKFZ !

Verbund: VerkäuferProdukt4NF

VerkäuferProduktKFZ

VerkNr	Produktname	KFZNr
1	Waschmaschine	M-E 515
1	Waschmaschine	M-X 333
1	Herd	M-E 515
1	Herd	M-X 333
1	Kühlschrank	M-E 515
1	Kühlschrank	M-X 333
2	Herd	S-H 654
2	Herd	K-J 123
2	Kühlschrank	S-H 654
2	Kühlschrank	K-J 123
3	Staubsauger	K-J 123

ProduktKFZ

Produktname	KFZNr
Waschmaschine	M-E 515
Herd	M-E 515
Herd	M-X 333
Herd	S-H 654
Herd	K-J 123
Kühlschrank	M-E 515
Kühlschrank	S-H 654
Staubsauger	K-J 123



Folgerung:

Wir erhalten VerkäuferProdukt4NF!
VerkäuferProdukt4NF besitzt Verbund-
abhängigkeit und ist nicht in 5. NF

Zusatzinfos zur 4. und 5. Normalform

- ▶ **Besitzt eine Relation nur nicht zusammengesetzte Schlüsselkandidaten, dann gilt:**
 - ▶ 3. NF (beide Versionen) = 4. NF = 5. NF
- ▶ **Eine Relation in 4. NF benötigt mindestens 3 Relationen zum Zerlegen in Relationen der 5. NF**
- ▶ **Relationen in mindestens der 3. NF, aber nicht in 4. oder 5. NF besitzen einen Schlüsselkandidaten mit mindestens drei Attributen!**

Zusammenfassung

- ▶ Ziel ist die dritte Normalform!
- ▶ Wir benötigen Wissen zu funktionaler Abhängigkeit!
- ▶ Wir erzeugen Relationen mit möglichst einfachen Schlüsselkandidaten
 - ▶ Somit benötigen wir kein Wissen zur 4. und 5. NF
 - ▶ Alle Relationen in 3. NF sind dann auch in 4. und 5. NF!